



---

**Towards a predictive first principles understanding of molecular adsorption on graphene 122099**

**Angelos Michaelides**  
**UNIVERSITY COLLEGE LONDON**

---

**10/05/2016**  
**Final Report**

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory  
AF Office Of Scientific Research (AFOSR)/ IOE  
Arlington, Virginia 22203  
Air Force Materiel Command

<b>REPORT DOCUMENTATION PAGE</b>				Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Executive Services, Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.</p>					
1. REPORT DATE (DD-MM-YYYY) 09-12-2016		2. REPORT TYPE Final		3. DATES COVERED (From - To) 30 Sep 2012 to 29 Sep 2016	
4. TITLE AND SUBTITLE Towards a predictive first principles understanding of molecular adsorption on graphene				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER FA8655-12-1-2099	
				5c. PROGRAM ELEMENT NUMBER 61102F	
6. AUTHOR(S) Angelos Michaelides				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) UNIVERSITY COLLEGE LONDON GOWER STREET LONDON, WC1E 6BT GB				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD Unit 4515 APO AE 09421-4515				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/AFOSR IOE	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-AFOSR-UK-TR-2016-0039	
12. DISTRIBUTION/AVAILABILITY STATEMENT A DISTRIBUTION UNLIMITED: PB Public Release					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT <p>The adsorption of atoms and molecules on surfaces is of central importance to an almost endless list of practical applications and everyday phenomena, such as corrosion, lubrication, friction, catalysis, coatings, and sensors to name but a few. In terms of understanding the properties of surfaces and the processes that occur on them, computer simulation techniques have been incredibly useful. However, even the most sophisticated quantum mechanical approaches struggle to make accurate predictions about the adsorption of molecules on surfaces, especially for the case of weakly interacting adsorbates. In this project we used and developed state-of-the-art quantum mechanical methods to make accurate predictions about the interaction strength and adsorption structure of molecules on a range of substrates and under confinement. This has led to new insight into how e.g. water binds to clay surfaces. An additional very positive outcome of this project is the development of an improved algorithm for quantum Monte Carlo simulations of complex materials, which provides both improved accuracy and improved computational efficiency.</p>					
15. SUBJECT TERMS EOARD, adsorption, graphene, computational materials science, density functional theory, ab initio methods					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 11	19a. NAME OF RESPONSIBLE PERSON FOLEY, JASON
a. REPORT  Unclassified	b. ABSTRACT  Unclassified	c. THIS PAGE  Unclassified			19b. TELEPHONE NUMBER (Include area code) 011-44-1895-616036

# Final Report: Grant No. FA8655-12-1-2099

## Towards a predictive first principles understanding of molecular adsorption on graphene

**Principal Investigator:** Prof. Angelos Michaelides, London Centre for Nanotechnology, Department of Physics and Astronomy, University College London, 17–19 Gordon Street, London WC1H 0AH, UK.

### 1. Introduction

The characterization of adsorption of atoms or molecules on different surfaces is a major challenge of materials modeling, and it is of central importance to an almost endless list of practical applications and everyday phenomena, such as corrosion, lubrication, friction, catalysis, coatings, crystal formation, and sensors to name but a few. Indeed, our perception of the physical properties of materials often depends sensitively on how adsorbates interact with them. In terms of understanding the properties of surfaces and the processes that occur on them, computer simulation techniques have been incredibly useful. Materials modeling *in silico*, made possible by the increasingly larger availability of compute power, is becoming ever more important for its use in industry and the positive outcomes for society. Nowadays, it is an essential part of science.

One of the foremost challenges in any computer simulation study of materials is to obtain an accurate description of how the atoms in the system interact with each other. In terms of adsorption at interfaces this is particularly challenging since many closed shell molecules (e.g. water, methane, benzene) interact weakly with their environment via the so-called “non-covalent” interactions, mainly due to van der Waals (vdW) dispersion forces. For small isolated systems there are established “quantum chemistry” based techniques for accounting for van der Waals dispersion forces. Examples include Møller-Plesset perturbation theory and coupled clusters. However, these approaches are almost exclusively limited to small system sizes (a few tens of atoms at most) and are not (yet) available for practical calculations of adsorption on periodic extended surfaces. For adsorption on surfaces, as well as in material modeling in general, density functional theory (DFT) remains the main workhorse, because of its computational affordability also in big systems. The problem, however, is that DFT with the most commonly used exchange-correlation functionals does not account for vdW dispersion forces. Recently our group in London (and other groups around the world) has been working on accurately including vdW dispersion forces within DFT. Indeed, the quest to accurately account for dispersion forces within DFT is one of the hottest topics in all of computational materials science at present. The development of these vdW inclusive functionals has been a breakthrough in the material science community, but still their accuracy is not sufficient to reliably evaluate important physical properties for a whole class of systems with weak non-covalent interactions, for example those involving the binding between water and most surfaces or in molecular crystals.

Non-covalent interactions are indeed a key ingredient in the structure, stability, and response properties of many molecular materials, which can be extremely sensitive to accuracy issues. While non-covalent interactions between small molecules in the gas phase are now relatively well

understood, much less certainty exists for binding properties of molecules and solids of increasing size and complexity. However, functional molecular complexes of interest in supramolecular and pharmaceutical applications are large and complex. Therefore, there is a pressing need to obtain reliable benchmarks for interaction energies between large molecules and in molecular crystals. In this context, quantum Monte Carlo (QMC) is the method of choice owing to its favorable computational scaling with system size. In several papers QMC was shown able to produce benchmark energies for non-covalent systems, reaching essentially the same accuracy as the coupled cluster with singles, doubles and perturbative triples (CCSD(T)) method, which is often regarded as the golden standard in quantum chemistry. The advantage of QMC over CCSD(T) is its applicability to large systems and the efficient use of petascale parallel supercomputers.

The initial aim of the project was to examine adsorption of small molecules on graphene with state of the art van der Waals functionals. Considerable progress was made in this direction, in particular for benzene on graphene. However, because of very exciting developments in QMC made as part of this project, results on a broader range of systems than just graphene have been obtained. Indeed, as a welcome byproduct of the project, we were able to spot a longstanding issue of size-inconsistency in QMC, and to devise a novel improved diffusion QMC (DMC) algorithm that solves this issue. The new method improves the accuracy of DMC and simultaneously is up to a hundred times faster than the older algorithm when applied to systems that interact through non-covalent interactions. Overall throughout the course of this project we have used a combination of DFT with advanced vdW inclusive functionals along with QMC and other quantum chemistry benchmarks. Where possible, our results have been cross validated against experiment and higher-level electronic structure theories.

Our close collaboration with Dario Alfè – also in the TYC at UCL and an expert in the application of QMC methods to adsorption on periodic surfaces – has been instrumental in the overall project. Moreover, the project has benefited from the collaboration with other members of Angelos Michaelides' group.

In the following, we will briefly highlight the main results obtained within this project. We will first report on the unexpected but extremely exciting improvement in the QMC method that came out of this project; work that was recently published in Phys. Rev. B. [1] Then, we will report about the results concerning the adsorption of molecules on a clay surface [2] and on graphene. It turned out that graphene was a particularly challenging system for QMC and by looking at a larger band-gap layered material (specifically a clay) considerably more progress could be made. Finally, we will report on results concerning the structural stability of water in recently discovered two-dimensional ice phases [3]. Each of these studies are either submitted for publication or about to be submitted.

## **2. Boosting the accuracy and speed of quantum Monte Carlo**

There are many materials and materials related problems for which DFT does not deliver the desired accuracy. For such problems explicitly correlated wave-function based approaches are needed, such as the approaches of quantum chemistry, QMC, and combinations thereof. In practice for condensed phase systems with more than a handful of atoms in the unit cell QMC remains the only feasible reference method, partly because of its favorable scaling with system size and the

fact that it can be used efficiently on massively parallel supercomputers. Indeed QMC, mostly within the DMC approach, is increasingly used to provide benchmark data for solids and to tackle interesting materials science problems that have been beyond the reach of DFT. DMC is also proving increasingly useful in exposing and helping to explain problems with DFT and as such helping to further the development of DFT.

DMC is in principle an exact technique to solve the imaginary time dependent Schrodinger equation. The discretization of time in practical implementations introduces a time-step ( $\tau$ ) error, with a computational cost proportional to  $1/\tau$ . Recently it was shown that for  $\text{CH}_4\text{-H}_2\text{O}$  clusters current implementations of DMC appear to be non size-consistent, i.e. the total energy of a system of  $M$  non-interacting molecules is not proportional to  $M$ . In Ref. [1] we show that this is a general problem, we identify its source, and propose a simple modification that solves it. Moreover, we observe that the time-step error in binding energy evaluations is mostly due to this size-consistency issue. Our proposed method also leads to remarkable speedups, by significantly increasing the accuracy of large  $\tau$  DMC evaluations.

All technical details can be found in ref. [1] and in its supporting information. The main methodological point is that the size-consistency issue in the old DMC scheme (originally devised by Umrigar, Nightingale and Runge in 1993, and hereafter referred as UNR) is present for any finite value of  $\tau$ , and disappears only in the limit of  $\tau \rightarrow 0$ . Since the main application area of DMC is the calculation of medium to large systems for which relatively small energy differences are computed but very small  $\tau$  cannot be afforded, this issue threatens the usefulness of DMC in material science. Our new algorithm is almost unaffected by this issue for a large range of possible values of  $\tau$ .

To quantify the size-consistency problem consider two systems A and B with energies  $E_A$  and  $E_B$ , and define  $E_{A,B}^{\text{separated}}$  the energy of the system with A and B at large enough distance from each other to have zero interaction. The quantity  $E_s = E_{A,B}^{\text{separated}} - (E_A + E_B)$  is therefore expected to be equal to zero, and if it is not it measures the size-consistency error. To compute the binding energy of the system where A and B are interacting and have a total energy  $E_{A,B}^{\text{bonded}}$  it is useful to define  $E_b = E_{A,B}^{\text{bonded}} - (E_A + E_B)$  and  $E_{bs} = E_{A,B}^{\text{bonded}} - E_s = E_{A,B}^{\text{separated}}$ . Here  $E_b$  may be affected by a size-consistency problem, and  $E_{bs}$  is not. To illustrate the problem we have selected here two representative examples.

Our first example is a system formed by a  $\text{CH}_4$  (A) and a  $\text{H}_2\text{O}$  (B) molecules  $E_{A,B}^{\text{separated}}$  is obtained for a C-O distance of 11.44 Angstrom. On the basis of CCSD(T) calculations we know that the residual interaction energy is  $< 0.1$  meV, negligible for our purposes.  $E_s$  is zero also for variational Monte Carlo (VMC), proving that the trial wavefunction of the dimer is effectively factorized. In Fig. 1 (top) we plot  $E_s$  computed with DMC as a function of  $\tau$ . For  $\tau \rightarrow 0$ ,  $E_s \rightarrow 0$  as expected, however, at a typical time-step  $\tau = 0.005$  a.u. the error is already  $\sim 20$  meV, which is about the same size of the binding energy of the dimer near the equilibrium distance, and grows to over 80 meV at  $\tau = 0.05$  a.u. In Fig. 1 (bottom) we show the binding energy of the molecule for a configuration near the equilibrium distance  $\text{CH}_4\text{-H}_2\text{O}(\text{bonded})$ , displayed in the inset of the figure. As expected from the large size-consistency problem highlighted above, the binding energy computed with  $E_b$  is wrong, and has a strong time-step dependence. Extrapolating to zero time-step using the whole range from 0.05 to 0.005 a.u. yields  $E_b = 11 \pm 7$  meV, and using only the range 0.02 to 0.005 a.u. a

value of  $E_b=21\pm 2$  meV, which is close to the benchmark energy  $E_b=24.5$  meV, obtained with CCSD(T). By contrast,  $E_{bs}$  is effectively time-step independent up to  $\tau=0.05$ , is in better agreement with the reference value, and removes the need for uncertain and arbitrary extrapolations. The UNR limiting procedure is too unstable above  $\tau=0.05$  and even at  $\tau=0.05$  we have not been able to obtain a very small statistical error due to a large number of instabilities.

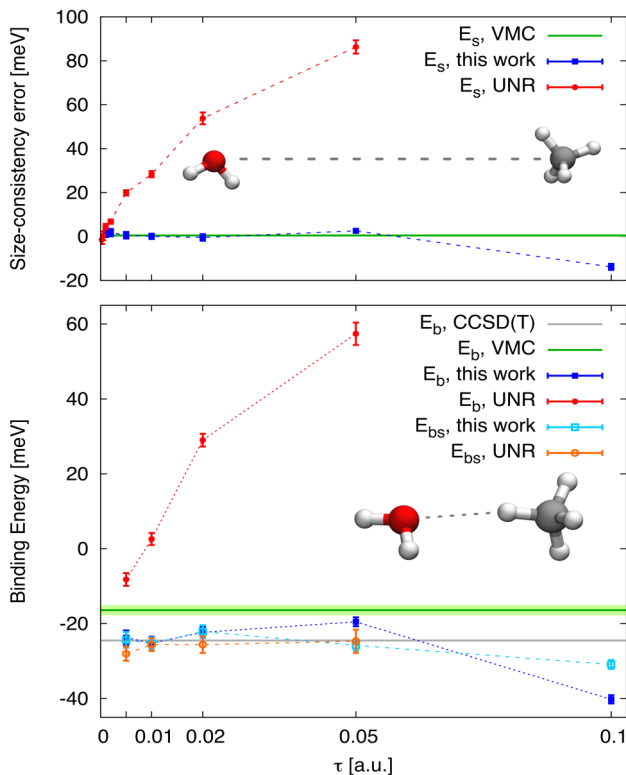


Figure 1: (Top) Size-consistency error  $E_s$  (see text) and (bottom) binding energy [using two different definitions,  $E_b$  and  $E_{bs}$  (see text)] for the  $\text{CH}_4\text{-H}_2\text{O}$  system, using DMC with the old algorithm (UNR) or the new one (this work). VMC and CCSD(T) evaluations are also shown. Error bars are one standard deviation. The insets show the structures of the complexes which have the molecules at large (top) and near the equilibrium (bottom) separation.

Although in some cases one could envisage to use  $E_{bs}$  to compute binding energies, it is much more desirable to be able to use  $E_b$  instead, particularly when one is concerned with the binding energy of more than just a dimer. For example, in the case of a cluster formed by a large number of molecules the construction of the system with all molecules far enough away from each other could be difficult, or even impossible, and alternative correction schemes would be required.

Binding energies computed with the new method are displayed in the bottom panel of Fig. 1, showing that  $E_{bs}$  has the same accuracy as that computed with the UNR branching factor, but now also  $E_b$  can be used reliably. The new method is stable also for  $\tau=0.1$  a.u., although at this very large value of time-step the binding energy starts to show non negligible errors. Note that in order to obtain a sufficiently high accuracy on  $E_b$  with the UNR branching factor, without relying on extrapolations, we would need to reduce the time-step at least down to  $\tau=0.0005$  a.u., which is two orders of magnitude smaller than what is required with our newly proposed method.

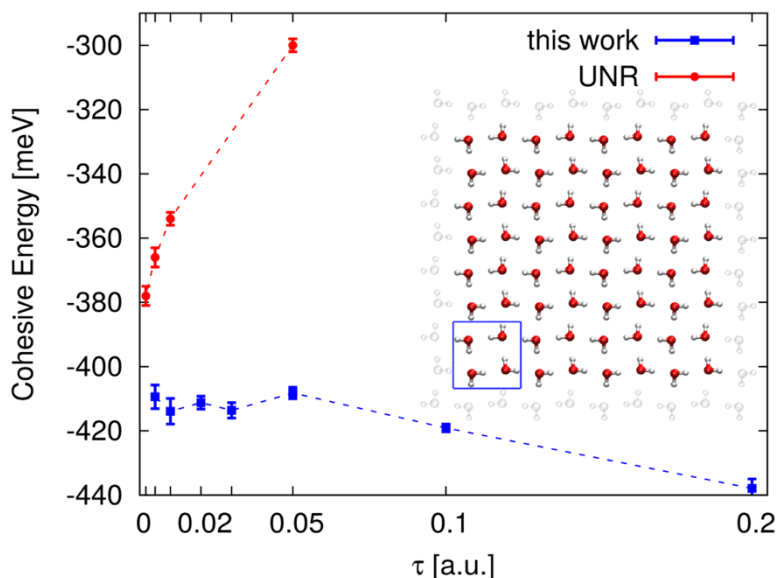


Figure 2: Cohesive energy of a two-dimensional periodic square ice system. The inset of the structure shows the simulated 64 molecule supercell as colored molecules, and the primitive unit cell in the blue square.

As a second example, we consider a square lattice ice system, a H-bonded 2D-periodic system subject of recent theoretical and experimental studies. The simulation cell comprises 64 water molecules. In Fig. 2 we show the cohesive energy as a function of time-step. The cohesive energy computed with the new limiting procedure is independent of time-step up to at least  $\tau=0.05$  a.u., while that computed with the UNR branching factor has errors that do not disappear even at the shortest time-step that we could afford ( $\tau=0.002$  a.u.). The non-linear trend of the UNR curve makes any  $\tau \rightarrow 0$  extrapolation unreliable, unless simulations with  $\tau < 0.001$  a.u. could be afforded. Given the size of this system this makes such calculation prohibitively expensive. Remarkably, the new method does not require to perform uncertain time-step extrapolations, and yields a speedup of around two orders of magnitude.

Another representative example, the so called buckyball catcher ( $C_{60}$ - $C_{60}H_{28}$ ) complex, is reported in Ref. [1] to highlight the outstanding performances of the new DMC algorithm in large supramolecular systems.

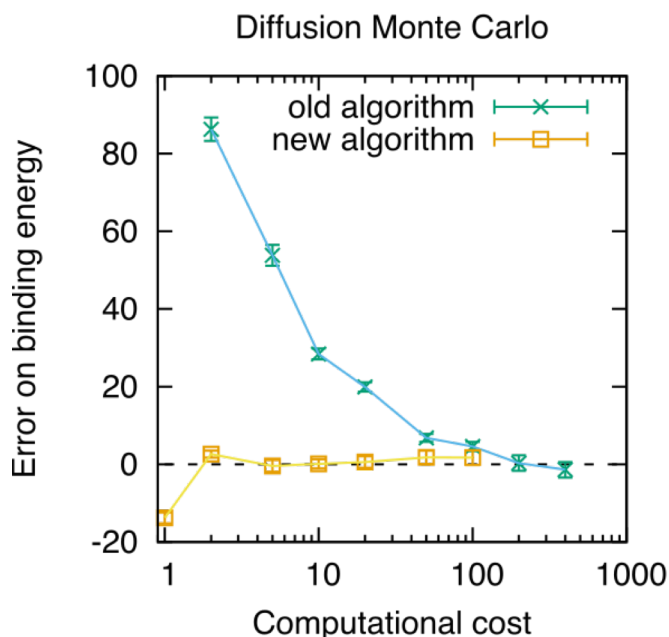


Figure 3: Size-consistency error as a function of the computational cost in DMC calculations obtained by using UNR (old algorithm) and present work (new algorithm) prescriptions. Computational cost is in arbitrary units.

In summary, we have proposed a procedure that reduces DMC time-step errors by a large factor, and restores size-consistency. Besides solving the size-consistency problem, speedups of two orders of magnitude are obtained (see Fig. 3) and the need for time-step extrapolations is removed. The improvement appears particularly promising for investigations on molecular materials and to discriminate between crystal polymorphs. Moreover, the recent emergence of QMC-based molecular dynamics, which until now have only been affordable within VMC, could now be in reach with the more accurate fixed-node DMC approach.

### 3. Towards a predictive first principles understanding of molecular adsorption on graphene

The many potential applications of graphene in materials science and technology, e.g. as a highly sensitive chemical sensor, is one of the key factors behind our interest in this particular material.

There is very little reliable data under clean conditions for how strongly molecules interact with graphene, since DFT fails to describe electron dispersion with standard approximations, indeed different state-of-the-art computer simulations methods disagree on how strong the interaction of the molecule with the surface is. This is illustrated in Fig. 4, which reports results on the performance of eleven different xc-functionals in describing the adsorption of benzene on graphene, a weakly interacting system dominated by dispersion. The improved performance of the dispersion schemes over the standard functionals popularly used in DFT is clear. The functionals that have better agreement with experiments, after applying the zero point energy (ZPE) correction, are vdW-DF (optB88-vdW and optB86b-vdW), and the  $C_6$  correction, PBE+vdW. These results clearly point at the necessity to apply dispersion developed functionals to describe weakly



interacting systems or dispersion corrections on the conventional semilocal and local xc-functionals.

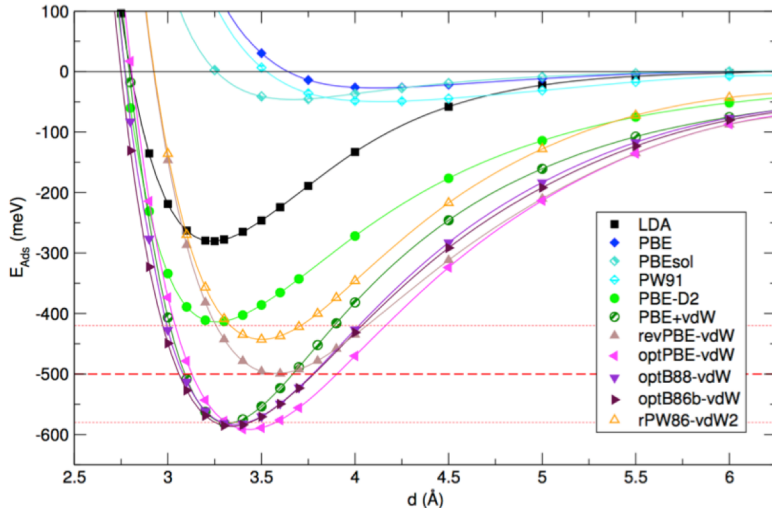


Figure 4: Binding curves of the benzene on graphene. The dashed lines represent the experimental reference for the binding energy and its error range.

However, the low precision of the experimental evaluation, plus the need to correctly account for ZPE corrections, makes a direct computational evaluation of the binding energy via a high level approach much more useful, especially in the perspective of helping with the development of the next generation of electronic structure method. DMC, after the methodological improvement described in the previous section [1], appears now as one of the most promising approaches to provide accurate benchmark values. Indeed, DMC can now be used to make accurate predictions about the interaction of molecules with surfaces which is essential for the development of better sensors and catalysts. Thus, we have performed DMC evaluations of the interaction energy of benzene and other molecules with graphene, and a paper is going to be submitted for publication soon.

#### 4. Toward Accurate Adsorption Energetics on Clay Surfaces

Clay minerals are ubiquitous in nature and the manner in which they interact with their surroundings has important industrial and environmental implications. Consequently, a molecular level understanding of the adsorption of molecules on clay surfaces is crucial. In this regard computer simulations play an important role, yet the accuracy of the most widely used empirical force field (FF) and DFT exchange-correlation functionals is unclear in general in adsorption systems dominated by weak interactions.

We have obtained results from QMC for water and methanol adsorption on the prototypical clay kaolinite. To the best of our knowledge, this is the first time QMC has been used to investigate adsorption at a complex, natural surface such as a clay. As well as being valuable in their own right, the QMC benchmarks obtained provide reference data against which the performance of cheaper DFT methods can be tested. Indeed using various DFT exchange-correlation functionals yields a very broad range of adsorption energies, and it is unclear *a priori* which evaluation is

better. QMC reveals that in the systems considered here it is essential to account for vdW dispersion forces since this alters both the absolute and relative adsorption energies of water and methanol, see Fig. 5.

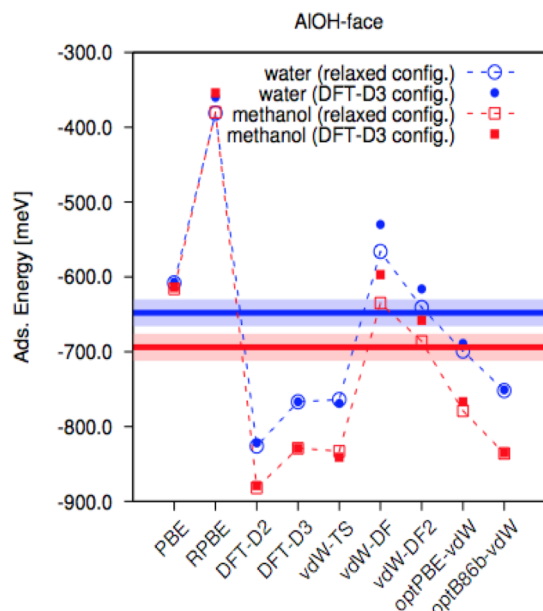


Figure 5: Adsorption of water (blue points) and methanol (red points) on the hydroxyl-terminated face of kaolinite, provided by different xc functionals and QMC calculations. Filled point and continuous lines represent the values for the configurations optimized using DFT-D3, and empty points and dashed lines correspond to relaxed geometries for each specific functional.

We have also shown, via FF simulations, that incorrect relative energies can lead to significant changes in the interfacial densities of water and methanol solutions at the kaolinite interface. Despite the clear improvements offered by the vdW-inclusive functionals, absolute adsorption energies are often overestimated, suggesting that the treatment of vdW forces in DFT is not yet a solved problem. A paper showing these results is submitted for publication [2].

## 5. Two-dimensional ice evaluated with Quantum Monte Carlo

Recent experiments on ice formed by water under nanoconfinement provide evidence for a new two-dimensional (2D) square ice phase. However, the interpretation of the initial experiments has been questioned and the stability of the square ice has become a matter of debate. Partially this is because the simulation approaches employed so far (force fields and density functional theory) struggle in accurately describing the very small energy differences between the relevant phases. The reliability of DFT and FF applied to ice is indeed a longstanding issue: the water phase diagram, both in bulk and 2D, is extremely challenging for any computational approach, because there are several competing phases with an energy range of only a few tens of milli-electronvolts per water. Thus, an accuracy of a few meV per water is needed in order to compare different ice polymorphs. This is often beyond the capabilities of DFT XC functionals and most FF approaches. Therefore, a study with a more accurate method is highly desirable.

Previous studies indicate that in 2D ice there are four structures, as reported in Fig. 6, competing

for stability at different pressure and temperature conditions. Thus, we have performed an accurate evaluation of the stabilities of these 2D ice structures using QMC. We find that at relatively high pressure square ice is indeed a stable phase, supporting the initial experimental claim. Moreover, at lower pressures a pentagonal ice phase (not yet observed experimentally) is stable, and at ambient pressure the pentagonal ice phase is quasi-degenerate with a hexagonal ice phase.

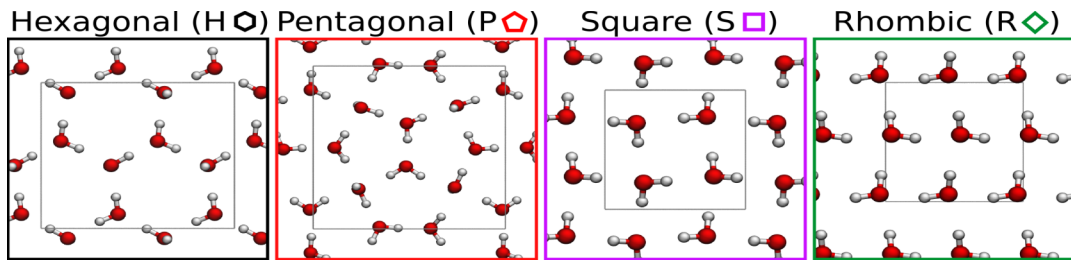


Figure 6: Monolayer 2D ice structures identified as stable or metastable in the literature: hexagonal (H), pentagonal (P), square (S), and rhombic (R). The gray box represents the primitive unit cell.

Our DMC results also allow us to evaluate the accuracy of various density functional theory exchange-correlation functionals and force field models, and in so doing we extend understanding of how such methods perform to challenging 2D structures presenting dangling hydrogen bonds. A paper showing these results is submitted for publication [3].

## 6. Summary and Conclusions

Overall this has been an incredibly interesting project scientifically. It is safe to say that it did not exactly go to plan. However, we think the outcomes achieved by far exceed what we aimed to do initially. In particular, the new QMC algorithm was a major unexpected bonus. It is very unlikely that we would have made this discovery if we were not able to move a post-doc on to the project, after the PhD student left. So we are very grateful to EOARD and Ty Pollak for showing such flexibility. The new QMC algorithm has great potential for application in a whole host of systems, ranging from adsorption problems, to van der Waals bonded solids, to molecular crystals. Aside from this the reference and benchmarking data for adsorption on graphene and on the clay, sheds new light on DFT and how various state of the art functionals perform.

## Relevant References

- [1] Zen, Andrea, Sandro Sorella, Michael J Gillan, Angelos Michaelides, and Dario Alfè. 2016. “Boosting the Accuracy and Speed of Quantum Monte Carlo: Size Consistency and Time Step.” *Physical Review B* 93: 241118(R)
- [2] Zen, Andrea; Roch, Loïc; Cox, Stephen; Hu, Xiao liang; Sorella, Sandro; Alfè, Dario; Michaelides, Angelos, Toward Accurate Adsorption Energetics on Clay Surfaces (submitted)
- [3] Ji Chen, Andrea Zen, Jan Gerit Brandenburg, Dario Alfe, and Angelos Michaelides, Two-dimensional ice evaluated with Quantum Monte Carlo (submitted)